

0590
0424

3



ENTERED

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/054,841

DATE: 04/25/2002 P.6
TIME: 12:04:27

Input Set : A:\19999YCA.txt

Output Set: N:\CRF3\04252002\J054841.raw

```

4 <110> APPLICANT: Chen, Fang
6 <120> TITLE OF INVENTION: DNA MOLECULES ENCODING HUMAN NUCLEAR
7   RECEPTOR PROTEINS
9 <130> FILE REFERENCE: 19999YCA
11 <140> CURRENT APPLICATION NUMBER: 10/054,841
12 <141> CURRENT FILING DATE: 2002-01-23
14 <150> PRIOR APPLICATION NUMBER: 09/487,379
15 <151> PRIOR FILING DATE: 2000-01-18
17 <150> PRIOR APPLICATION NUMBER: 09/141,000
18 <151> PRIOR FILING DATE: 1998-08-26
20 <150> PRIOR APPLICATION NUMBER: 60/078,633
21 <151> PRIOR FILING DATE: 1998-03-19
23 <150> PRIOR APPLICATION NUMBER: 60/062,902
24 <151> PRIOR FILING DATE: 1997-10-21
26 <150> PRIOR APPLICATION NUMBER: 60/057,090
27 <151> PRIOR FILING DATE: 1997-08-27
29 <160> NUMBER OF SEQ ID NOS: 30
31 <170> SOFTWARE: FastSEQ for Windows Version 4.0
33 <210> SEQ ID NO: 1
34 <211> LENGTH: 2807
35 <212> TYPE: DNA
36 <213> ORGANISM: Human
38 <400> SEQUENCE: 1
39 gaatatgatg accctaattgc aacaatatct aacatactat ccgagcttcg gtcatttgga      60
40 agaactgcag attttcctcc ttcaaaatta agtcagggtt atggagaaca tgtatgctat      120
41 gttcttgatt gcttcgctga agaagcattg aaatatattg gtttcacctg gaaaaggcca      180
42 atataccagc tagaagaatt agaagaagaa agcgttcgag aagatgatgc agaattaaca      240
43 ttaaataaag tggatgaaga atttgtggaa gaagagacag ataatgaaga aaactttatt      300
44 gatctcaacg ttttaaaggc ccagacatat cacttgata tgaacgagac tgccaaacaa      360
45 gaagatattt tggaatccac aacagatgct gcagaatgga gcctagaagt ggaacgtgta      420
46 ctaccgcaac tgaaagtcac gattaggact gacaataagg attggagaat ccatgttgac      480
47 caaatgcacc agcacagaag tggaattgaa tctgtcttaa aggagaccaa gggatttttg      540
48 gacaaactcc ataatgaaat tactaggact ttggaaaaga tcagcagccg agaaaagtac      600
49 atcaacaatc agccgggagc ccatggagca ctgtcctcag agatgcgcag gttaggctca      660
50 ctgtctaggc caggcccacc ttagtcaactg tggactggca atggaagctc ttcttgga      720
51 cacctgccct agccctcacc ctgggggtgga agagaaatga gcttggttg caactcagac      780
52 cattccacgg aggcacccctc cccttcctct ggctggtgaa taaaagtttc ctgaggtcaa      840
53 ggaacttcct ttccctgcc aaatggtgtc cagaactttg aggccagagg tgatccagtg      900
54 atttgggagc tgcaggtcac acaggctgct cagagggctg ctgaacagga tgtcctcgga      960
55 cgacaggcac ctggggtcca gctgcggctc ctcatcaag actgagccgt ccagcccgtc      1020
56 ctggggcata gatgccctca gccaccacag cccagtggtc tcgtccgacg ccagcggcgg      1080
57 ctttggcctg gccctgggca cccacgcca cggctctggac tcgccacca tgtttgcagg      1140
58 cgccgggctg ggaggcacc catgccgcaa gagctacgag gactgtgcca gcggcatcat      1200

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/054,841

DATE: 04/25/2002

TIME: 12:04:27

Input Set : A:\19999YCA.txt

Output Set: N:\CRF3\04252002\J054841.raw

```

59 ggaggactcg gccatcaagt gcgagtacat gctcaacgcc atccccaagc gcctgtgcct 1260
60 cgtgtgcggg gacattgcct ctggctacca ctacggcggtg gcctcctgcg aggcttgcaa 1320
61 ggcccttcttc aagaggacta tccaagggaa cattgagtac agctgcccgg ccaccaacga 1380
62 gtgcgagatc accaaacgga ggcgcaagtc ctgccaggcc tgccgcttca tgaaatgcct 1440
63 caaagtgggg atgctgaagg aaggtgtgctg ccttgatcga gtgcgtggag gccgtcagaa 1500
64 atacaagcga cggctggact cagagagcag cccatacctg agcttacaaa tttctccacc 1560
65 tgctaaaaag ccattgacca agattgtctc atacctactg gtggctgagc cggacaagct 1620
66 ctatgccatg cctccccctg gtatgcctga gggggacatc aaggccctga ccactctctg 1680
67 tgacctggca gccgagagc ttgtggtcat cattggctgg gccaaagcaca tcccaggctt 1740
68 ctcaagcctc tccctggggg accagatgag cctgctgcag agtgccctgga tggaaatcct 1800
69 catcctgggc atcgtgtacc gctcgtgcc ctacgacgac aagctggtgt acgctgagga 1860
70 ctacatcatg gatgaggagc actcccgctc cgcggggctg ctggagctct accgggccat 1920
71 cctgcagctg gtacgcaggt acaagaagct caaggtggag aaggaggagt ttgtgacgct 1980
72 caaggccctg gccctcgcca actccgattc catgtacatc gaggatctag aggctgtcca 2040
73 gaagctgcag gacctgctgc acgaggcact gcaggactac gagctgagcc agcgccatga 2100
74 ggagccctgg aggacgggca agctgctgct gacactgccg ctgctgcccg agacggccgc 2160
75 caaggcgtg cagcacttct atagcgtcaa actgcagggc aaagtgccca tgcacaaact 2220
76 ctctcctgag atgctggagg ccaaggcctg ggccagggtg gactcccttc aggagtggag 2280
77 gccactggag caagtgcctc ctccccctca ccgagccacc aagaggcagc atgtgcattt 2340
78 cctaactccc ttgccccctc ccccatctgt ggctgggtg ggcactgctc aggctggata 2400
79 ccacctggag gttttccttc cgcagagggc aggttgccca agagcagctt agaggatctc 2460
80 ccaaggatga aagaatgtca agccatgatg gaaaatgcc cttccaatca gctgccttca 2520
81 caagcagggg tcagagcaac tccccgggga tcccgaatcc acgcccctct agtccaaccc 2580
82 ccttcaatga gagaggcagg cagatctcac ccagcactag gacaccagga ggccagggaa 2640
83 agcatctctg gctcaccatg taacatctgg cttggagcaa gtgggtgttc tgcacaccag 2700
84 gcagctgcac ctactggat ctagtgttgc tgcagtgac ctacttcag agcccctcta 2760
85 gcagagtggg gcggaagtcc tgatggttgg tgtccatgag gtggaag 2807

```

87 <210> SEQ ID NO: 2

88 <211> LENGTH: 500

89 <212> TYPE: PRT

90 <213> ORGANISM: Human

92 <400> SEQUENCE: 2

```

93 Met Ser Ser Asp Asp Arg His Leu Gly Ser Ser Cys Gly Ser Phe Ile
94 1 5 10 15
95 Lys Thr Glu Pro Ser Ser Pro Ser Ser Gly Ile Asp Ala Leu Ser His
96 20 25 30
97 His Ser Pro Ser Gly Ser Ser Asp Ala Ser Gly Gly Phe Gly Leu Ala
98 35 40 45
99 Leu Gly Thr His Ala Asn Gly Leu Asp Ser Pro Pro Met Phe Ala Gly
100 50 55 60
101 Ala Gly Leu Gly Gly Thr Pro Cys Arg Lys Ser Tyr Glu Asp Cys Ala
102 65 70 75 80
103 Ser Gly Ile Met Glu Asp Ser Ala Ile Lys Cys Glu Tyr Met Leu Asn
104 85 90 95
105 Ala Ile Pro Lys Arg Leu Cys Leu Val Cys Gly Asp Ile Ala Ser Gly
106 100 105 110
107 Tyr His Tyr Gly Val Ala Ser Cys Glu Ala Cys Lys Ala Phe Phe Lys
108 115 120 125
109 Arg Thr Ile Gln Gly Asn Ile Glu Tyr Ser Cys Pro Ala Thr Asn Glu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/054,841

DATE: 04/25/2002

TIME: 12:04:27

Input Set : A:\19999YCA.txt

Output Set: N:\CRF3\04252002\J054841.raw

```

110      130      135      140
111 Cys Glu Ile Thr Lys Arg Arg Arg Lys Ser Cys Gln Ala Cys Arg Phe
112 145      150      155      160
113 Met Lys Cys Leu Lys Val Gly Met Leu Lys Glu Gly Val Arg Leu Asp
114      165      170      175
115 Arg Val Arg Gly Gly Arg Gln Lys Tyr Lys Arg Arg Leu Asp Ser Glu
116      180      185      190
117 Ser Ser Pro Tyr Leu Ser Leu Gln Ile Ser Pro Pro Ala Lys Lys Pro
118      195      200      205
119 Leu Thr Lys Ile Val Ser Tyr Leu Leu Val Ala Glu Pro Asp Lys Leu
120      210      215      220
121 Tyr Ala Met Pro Pro Pro Gly Met Pro Glu Gly Asp Ile Lys Ala Leu
122 225      230      235      240
123 Thr Thr Leu Cys Asp Leu Ala Asp Arg Glu Leu Val Val Ile Ile Gly
124      245      250      255
125 Trp Ala Lys His Ile Pro Gly Phe Ser Ser Leu Ser Leu Gly Asp Gln
126      260      265      270
127 Met Ser Leu Leu Gln Ser Ala Trp Met Glu Ile Leu Ile Leu Gly Ile
128      275      280      285
129 Val Tyr Arg Ser Leu Pro Tyr Asp Asp Lys Leu Val Tyr Ala Glu Asp
130      290      295      300
131 Tyr Ile Met Asp Glu Glu His Ser Arg Leu Ala Gly Leu Leu Glu Leu
132 305      310      315      320
133 Tyr Arg Ala Ile Leu Gln Leu Val Arg Arg Tyr Lys Lys Leu Lys Val
134      325      330      335
135 Glu Lys Glu Glu Phe Val Thr Leu Lys Ala Leu Ala Leu Ala Asn Ser
136      340      345      350
137 Asp Ser Met Tyr Ile Glu Asp Leu Glu Ala Val Gln Lys Leu Gln Asp
138      355      360      365
139 Leu Leu His Glu Ala Leu Gln Asp Tyr Glu Leu Ser Gln Arg His Glu
140      370      375      380
141 Glu Pro Trp Arg Thr Gly Lys Leu Leu Leu Thr Leu Pro Leu Leu Arg
142 385      390      395      400
143 Gln Thr Ala Ala Lys Ala Val Gln His Phe Tyr Ser Val Lys Leu Gln
144      405      410      415
145 Gly Lys Val Pro Met His Lys Leu Phe Leu Glu Met Leu Glu Ala Lys
146      420      425      430
147 Ala Trp Ala Arg Ala Asp Ser Leu Gln Glu Trp Arg Pro Leu Glu Gln
148      435      440      445
149 Val Pro Ser Pro Leu His Arg Ala Thr Lys Arg Gln His Val His Phe
150      450      455      460
151 Leu Thr Pro Leu Pro Pro Pro Pro Ser Val Ala Trp Val Gly Thr Ala
152 465      470      475      480
153 Gln Ala Gly Tyr His Leu Glu Val Phe Leu Pro Gln Arg Ala Gly Trp
154      485      490      495
155 Pro Arg Ala Ala
156      500
158 <210> SEQ ID NO: 3
159 <211> LENGTH: 2985

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/054,841

DATE: 04/25/2002

TIME: 12:04:27

Input Set : A:\19999YCA.txt

Output Set: N:\CRF3\04252002\J054841.raw

```

160 <212> TYPE: DNA
161 <213> ORGANISM: Human
163 <400> SEQUENCE: 3
164 gcggggccgc agtgtggtgg aattcggttt gtcactagga gaacatttgt gtttaattgca      60
165 ctgtgctctg tcaaggaaac tttgatattat agctgggggtg cacaaataat ggttgccggt      120
166 cgcacatgga ttcggtagaa ctttgcccttc ctgaatcttt ttccctgcac tacgaggaaag      180
167 agcttctctg cagaatgtca aacaaagatc gacacattga ttccagctgt tcgtccttca      240
168 tcaagacgga accttccagc ccagcctccc tgacggacag cgtcaaccac cacagccctg      300
169 gtggctcttc agacgccagt gggagctaca gttcaacatc gaatggccat cagaacggac      360
170 ttgactcgcc acctctctac ctttctgctc ctatcctggg aggtagtggg cctgtcagga      420
171 aactgtatga tgactgctcc agcaccattg ttgaagatcc ccagaccaag tgtgaataca      480
172 tgctcaactc gatgcccaag agactgtggt tagtgtgtgg tgacatcgct tctgggtacc      540
173 actatggggg agcatcatgt gaagcctgca aggcattctt caagaggaca attcaaggca      600
174 atatagaata cagctgccct gccacgaatg aatgtgaaat cacaaagcgc agacgtaaat      660
175 cctgccaggc ttgccgcttc atgaagtgtt taaaagtggg catgctgaaa gaaggggtgc      720
176 gtcttgacag agtacgtgga ggtcggcaga agtacaagcg caggatagat gcggagaaca      780
177 gccatacct gaacctcag ctggttcagc cagccaaaaa gccatataac aagattgtct      840
178 cacatttggt ggtggctgaa ccggagaaga tctatgccat gcctgaccct actgtccccg      900
179 acagtgcacat caaagccctc actacactgt gtgacttggc cgaccgagag ttggtggtta      960
180 tcattggatg ggcgaagcat attccaggct tctccacgct gtccctggcg gaccagatga     1020
181 gccttctgca gagtgcctgg atggaaattt tgatccttgg tgcgtatac cggctctctt     1080
182 catttgagga tgaacttgct tatgcagacg attatataat ggacgaagac cagtccaaat     1140
183 tagcaggcct tcttgatcta aataatgcta tcttgacgct ggtaaagaaa tacaagagca     1200
184 tgaagctgga aaaagaagaa tttgtcacc ctaaagctat agctcttgct aattcagact     1260
185 ccattgcacat agaagatggt gaagccgttc agaagcttca ggaatgtctta catgaagcgc     1320
186 tgcaggatta tgaagctggc cagcacatgg aagaccctcg tcgagctggc aagatgctga     1380
187 tgacactgcc actcctgagg cagacctcta ccaaggccgt gcagcatttc tacaacatca     1440
188 aactagaagg caaagtccca atgcacaaac tttttttgga aatgttggag gccaaagtct     1500
189 gactaaaagc tccctgggcc ttcccatcct tcatgttgaa aaagggaaaa taaacccaag     1560
190 agtgatgtcg aagaaactta gagttagttt aacaacatca aaaatcaaca gactgcactg     1620
191 ataatttagc agcaagacta tgaagcagct ttcagattcc tccataggtt cctgatgagt     1680
192 tctttctact ttctccatca tcttctttcc tcttcttcc cactttctc tttctcttta     1740
193 tttttctctc ttttctctt tcaacctcct tatttctttg cttctttcat tctagtctcc     1800
194 cattctcctt tatttcttcc ccgtctgctt gccttctttc ttttctttac ctactctcat     1860
195 tctctctctt tctcatcctt cccctttttt ctaaatttga aatagcttta gtttaaaaaa     1920
196 aaaaatcctc ccttccccct ttcccttccc tttcttctc ttttcccttt ccttttccct     1980
197 ttcccttctc ttccctctga ccttctttcc atcttctttt ttcttcttcc tgcgtctgaa     2040
198 ctttttaaaag aggtctctaa ctgaagagag atggaagcca gccctgccaaggatggaga     2100
199 tccataatat ggaatgcagt gaacttattg tgaaccatac cgtccccaat gactaaggaa     2160
200 tcaaagagag agaaccaacg ttccctaaaag tacagtgcac catatacaaa ttgactgagt     2220
201 gcagtattag atttcatggg agcagcctct aattagacaa ctttaagcaac gttgcatcgg     2280
202 ctgcttctta tcattgcttt tccatctaga tcagttacag ccatttgatt ccttaattgt     2340
203 tttttcaagt cttccaggta tttgttagtt tagctactat gtaacttttt cagggaatag     2400
204 ttttaagcttt attcattcat gcaatactaa agagaaataa gaatactgca attttgtgct     2460
205 ggctttgaac aattacgaac aataatgaag gacaaatgaa tcctgaagga agatttttaa     2520
206 aaatgttttg tttctcttta caaatggaga tttttttgta ccagctttac cacttttcag     2580
207 ccattttatta atatgggaat ttaacttact taagcaatag ttgaagggaa ggtgcattat     2640
208 atcacggatg caatttatgt tgtgtgccag tctggtccca aacatcaatt tcttaacatg     2700
209 agctccagtt tacctaaatg ttcactgaca caaaggatga gattacacct acagtgactc     2760

```

RAW SEQUENCE LISTING

DATE: 04/25/2002

PATENT APPLICATION: US/10/054,841

TIME: 12:04:27

Input Set : A:\19999YCA.txt

Output Set: N:\CRF3\04252002\J054841.raw

```

210 tgagtagtca catatataag cactgcacat gagatataga tccgtagaat tgtcaggagt      2820
211 gcacctctct acttgggagg tacaattgcc atatgatttc tagctgccat ggtgggttagg      2880
212 aatgtgatac tgctgtttg caaagtcaca gaccttgccct cagaaggagc tgtgagccag      2940
213 tattcattta agagaattcc accacactgg cggcccgcgc ttgat                        2985
215 <210> SEQ ID NO: 4
216 <211> LENGTH: 458
217 <212> TYPE: PRT
218 <213> ORGANISM: Human
220 <400> SEQUENCE: 4
221 Met Asp Ser Val Glu Leu Cys Leu Pro Glu Ser Phe Ser Leu His Tyr
222 1 5 10 15
223 Glu Glu Glu Leu Leu Cys Arg Met Ser Asn Lys Asp Arg His Ile Asp
224 20 25 30
225 Ser Ser Cys Ser Ser Phe Ile Lys Thr Glu Pro Ser Ser Pro Ala Ser
226 35 40 45
227 Leu Thr Asp Ser Val Asn His Ser Pro Gly Gly Ser Ser Asp Ala
228 50 55 60
229 Ser Gly Ser Tyr Ser Ser Thr Met Asn Gly His Gln Asn Gly Leu Asp
230 65 70 75 80
231 Ser Pro Pro Leu Tyr Pro Ser Ala Pro Ile Leu Gly Gly Ser Gly Pro
232 85 90 95
233 Val Arg Lys Leu Tyr Asp Asp Cys Ser Ser Thr Ile Val Glu Asp Pro
234 100 105 110
235 Gln Thr Lys Cys Glu Tyr Met Leu Asn Ser Met Pro Lys Arg Leu Cys
236 115 120 125
237 Leu Val Cys Gly Asp Ile Ala Ser Gly Tyr His Tyr Gly Val Ala Ser
238 130 135 140
239 Cys Glu Ala Cys Lys Ala Phe Phe Lys Arg Thr Ile Gln Gly Asn Ile
240 145 150 155 160
241 Glu Tyr Ser Cys Pro Ala Thr Asn Glu Cys Glu Ile Thr Lys Arg Arg
242 165 170 175
243 Arg Lys Ser Cys Gln Ala Cys Arg Phe Met Lys Cys Leu Lys Val Gly
244 180 185 190
245 Met Leu Lys Glu Gly Val Arg Leu Asp Arg Val Arg Gly Gly Arg Gln
246 195 200 205
247 Lys Tyr Lys Arg Arg Ile Asp Ala Glu Asn Ser Pro Tyr Leu Asn Pro
248 210 215 220
249 Gln Leu Val Gln Pro Ala Lys Lys Pro Tyr Asn Lys Ile Val Ser His
250 225 230 235 240
251 Leu Leu Val Ala Glu Pro Glu Lys Ile Tyr Ala Met Pro Asp Pro Thr
252 245 250 255
253 Val Pro Asp Ser Asp Ile Lys Ala Leu Thr Thr Leu Cys Asp Leu Ala
254 260 265 270
255 Asp Arg Glu Leu Val Val Ile Ile Gly Trp Ala Lys His Ile Pro Gly
256 275 280 285
257 Phe Ser Thr Leu Ser Leu Ala Asp Gln Met Ser Leu Leu Gln Ser Ala
258 290 295 300
259 Trp Met Glu Ile Leu Ile Leu Gly Val Val Tyr Arg Ser Leu Ser Phe
260 305 310 315 320

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/054,841

DATE: 04/25/2002
TIME: 12:04:28

Input Set : A:\19999YCA.txt
Output Set: N:\CRF3\04252002\J054841.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 279,294
Seq#:8; N Pos. 2,3,4,5,6,7,8,9,10,11,12,13,18,19,37,42,401,471,521,532,534
Seq#:8; N Pos. 550,556,557,558,559,560,561,562,563,564,565,566,567,568,569
Seq#:8; N Pos. 570,571,572,573,574,575,576,577,578,579,580,581,582,583,584
Seq#:8; N Pos. 585,586,587,588,589,590,591,592,593,594,595,596,597,598,599
Seq#:8; N Pos. 600,601,602,603,604,605,606,607,608,609,610,611,612,613,614
Seq#:8; N Pos. 615,616,617,618,619,620

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/054,841

DATE: 04/25/2002

TIME: 12:04:28

Input Set : A:\19999YCA.txt

Output Set: N:\CRF3\04252002\J054841.raw

L:413 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:240
L:428 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0
L:434 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:360
L:435 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:420
L:436 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:480
L:437 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:540
L:438 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:600